

Using PhenoGen To Investigate a Gene

Outline

- Examine BN-Lx/SHR brain RNA-Seq results (reconstruction; splice junction reads; coverage)
- Examine brain RNA-Seq results from other inbred strains
- Gene Details
- Gene eQTL
- Expression Data
- miRNA Targeting Gene (multiMir)
- Co-expression modules (WGCNA)

Examine BN-Lx/SHR Brain RNA-
Seq Results

<http://phenogen.org>

Select
'Genome/Transcriptome
Data Browser'

The screenshot shows the PhenoGen Informatics homepage. At the top, there is a banner for a "Summer of Rat Genomics and Genetics Workshop - Aug 2nd-3rd, 2019 - Colorado - NIDA Center for Omics, Systems Genetics and the Addictome" with links for "More Info" and "Registration". Below the banner is a dark blue header with a "Google Custom Search" bar and a search icon. The main navigation menu consists of several items: "Overview" (with a dropdown arrow), "Genome / Transcriptome Data Browser" (which is circled in red and has a red arrow pointing to it from the text above), "Available Data Downloads", "Gene List Analysis Tools" (with a dropdown arrow), "About" (with a dropdown arrow), "Help" (with a dropdown arrow), "View Previous Microarray Analysis" (with a dropdown arrow), and "Login / Register" (with a dropdown arrow). Below the header, a large white section welcomes users with the text "Welcome to PhenoGen Informatics" and "The site for quantitative genetics of the transcriptome." In the bottom left corner of this section, there is a note: "Hover over or click on nodes in the graph below to see the tools/data available on the site".

Input Gene of Interest (e.g., Mef2c)

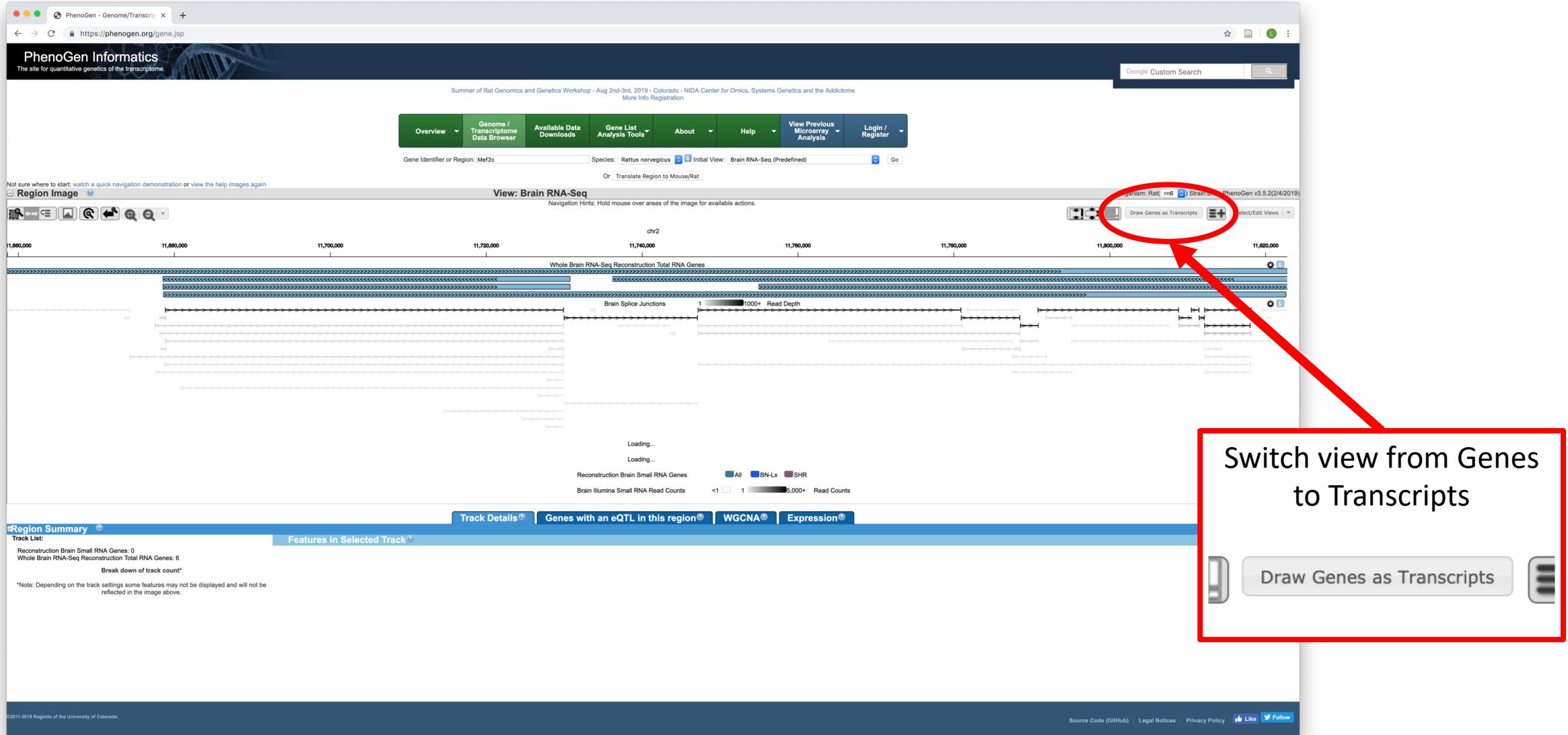
Gene Identifier or Region: Species: ⓘ Initial View: ⏺ Go

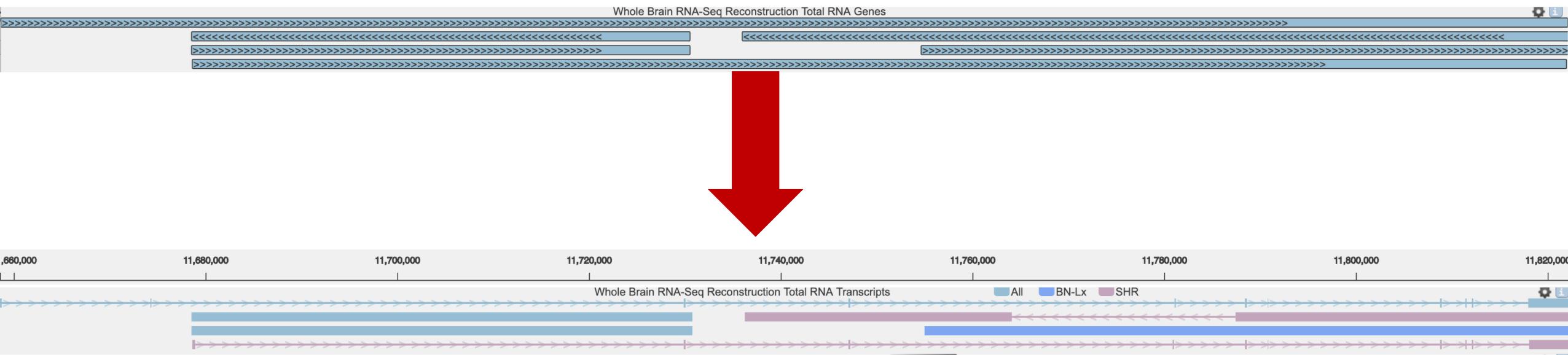
Or [Translate Region to Mouse/Rat](#)

Demonstrations
Custom View/Custom Track Demonstration

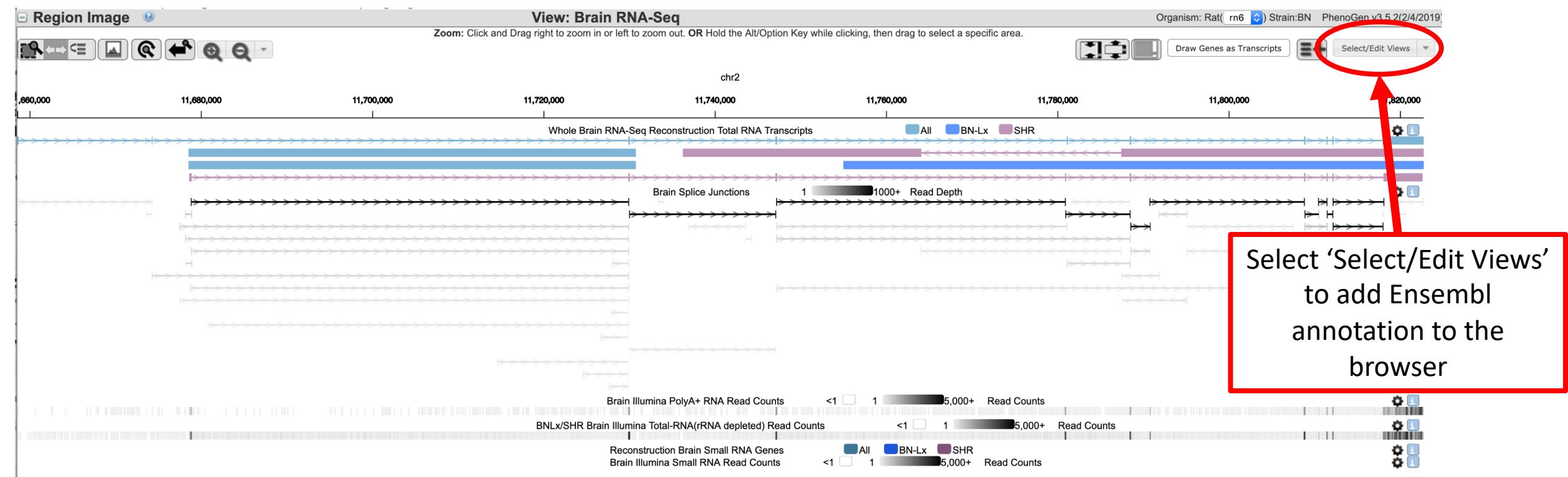
Enter Gene Symbol of Interest

Select 'Brain RNA-Seq (Predefined)' from the drop-down menu for initial view





Add Ensembl Annotation To Genome Viewer



1. Click on 'View/Edit Track List' Tab

2. Click on Icon to add track

The screenshot shows the 'Select/Edit Views' dialog box. At the top, there are three icons: a blue plus sign, a blue document with a double arrow, and a blue folder. Below them is a dropdown menu set to 'All Views'. To the right, a message says 'Click on a view to select it and view preview/details.' and a 'Apply View' button is visible. The main area lists several predefined views: 'Genome (Predefined) (8 tracks)', 'Transcriptome (Predefined) (12 tracks)', 'Genome/Transcriptome (Predefined) (15 tracks)', 'Liver RNA-Seq (Predefined) (7 tracks) (Rat Only)', 'Brain RNA-Seq (Predefined) (7 tracks)', 'Liver Microarray/RNA-Seq (Predefined) (8 tracks) (Rat Only)', 'Brain Microarray/RNA-Seq (Predefined) (8 tracks)', and 'Heart RNA-Seq (Predefined) (7 tracks) (Rat Only)'. Below this is a table with two tabs at the top: 'Description/Preview' (selected) and 'View/Edit Track List'. The 'View/Edit Track List' tab is circled in red. The table has columns for Order, Track Name, Organism, and Edit. Seven rows are listed, each with a small icon and edit controls. The first row is 'Reference Genomic Sequence'.

Order	Track Name	Organism	Edit
1	Reference Genomic Sequence		↑ ↓ ⚙ ✘
2	Brain Total-RNA Transcriptome		↑ ↓ ⚙ ✘
3	Brain Splice Junction Support	Rat only	↑ ↓ ⚙ ✘
4	Brain Illumina PolyA+ RNA	Rat only	↑ ↓ ⚙ ✘
5	Brain Illumina rRNA-depleted Total-RNA	Rat only	↑ ↓ ⚙ ✘
6	Brain Transcriptome Small RNA	Rat only	↑ ↓ ⚙ ✘
7	Brain Illumina small RNA	Rat only	↑ ↓ ⚙ ✘

Apply View

Select a Track to add to Brain RNA-Seq X

⚠ Sign in to see tracks not saved to this computer. (Any tracks created will only be saved locally) i

Add Track ?

Available Tracks:

View types: All Tracks ▼

Search:

Track Name	Organism	Genome / Transcriptome	Category	Public / Custom
Ensembl Protein Coding Genes i		Genome	Annotation	Public
Ensembl Long Non-Coding Genes i		Genome	Annotation	Public
Ensembl Small RNA Genes i		Genome	Annotation	Public
Ref Seq Genes i		Genome	Annotation	Public
UCSC Repeat Masker i		Genome	Annotation	Public
Affymetrix Exon 1.0ST Probes i	Rat	Transcriptome	Microarray	Public
bQTLs i		Genome	QTL	Public
Liver + Strand Total-RNA (BN-Lx/SHR) i	Rat	Transcriptome	RNA-Seq Read Depth	Public
Liver - Strand Total-RNA (BN-Lx/SHR) i	Rat	Transcriptome	RNA-Seq Read Depth	Public
Heart + Strand Total-RNA (BN-Lx/SHR) i	Rat	Transcriptome	RNA-Seq Read Depth	Public
Heart - Strand Total-RNA (BN-Lx/SHR) i	Rat	Transcriptome	RNA-Seq Read Depth	Public
Liver + Strand Total-RNA (Cop) i	Rat	Transcriptome	RNA-Seq Read Depth	Public
Liver - Strand Total-RNA (Cop) i	Rat	Transcriptome	RNA-Seq Read Depth	Public
Liver + Strand Total-RNA (Dark-Agouti) i	Rat	Transcriptome	RNA-Seq Read Depth	Public

1. Select Track of Interest



Ensembl Protein Coding Genes i

Add Track

2. Click 'Add Track'



Select/Edit Views

 Sign in to see views/tracks not created on this computer. [i](#)



View types: All Views 

Apply View

Select a view below (click apply to display the view and return to the browser):

- Genome (Predefined) (8 tracks)
- Transcriptome (Predefined) (12 tracks)
- Genome/Transcriptome (Predefined) (15 tracks)
- Liver RNA-Seq (Predefined) (7 tracks) (Rat Only)
- Brain RNA-Seq (Predefined) (7 tracks)**
- Liver Microarray/RNA-Seq (Predefined) (8 tracks) (Rat Only)
- Brain Microarray/RNA-Seq (Predefined) (8 tracks)
- Heart RNA-Seq (Predefined) (7 tracks) (Rat Only)

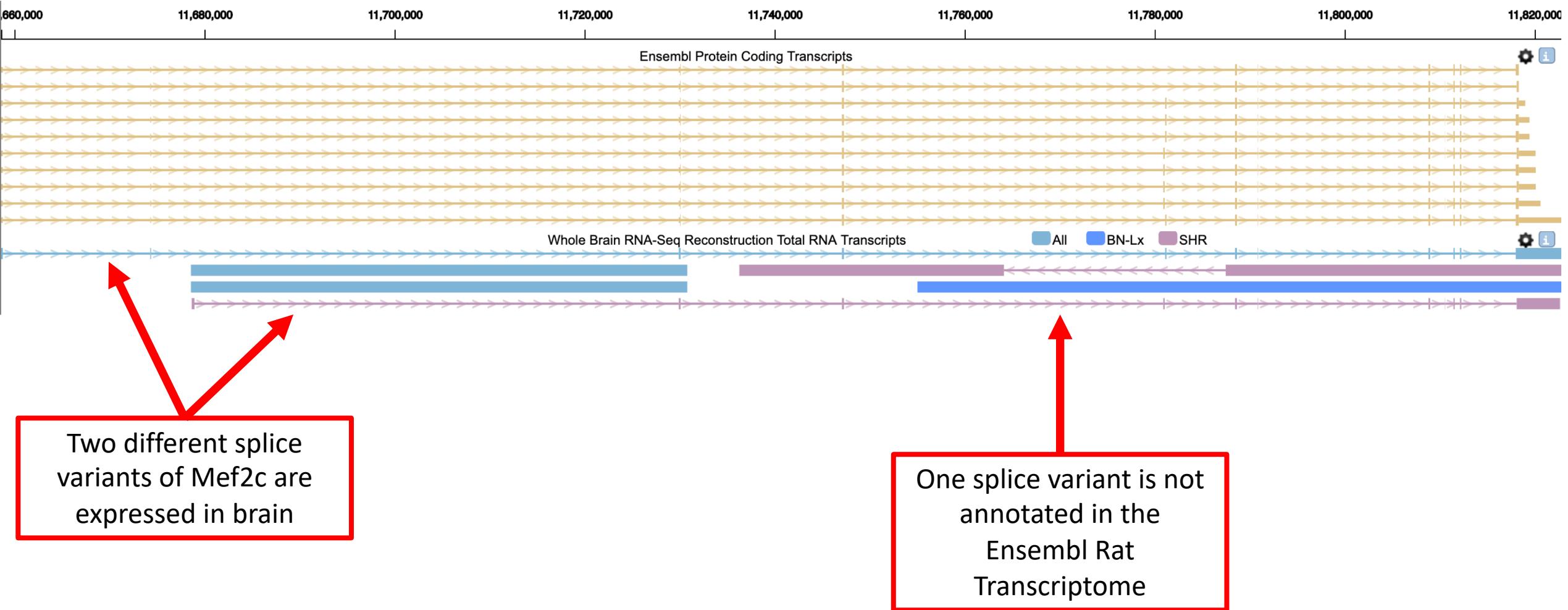
Description/Preview

View/Edit Track List

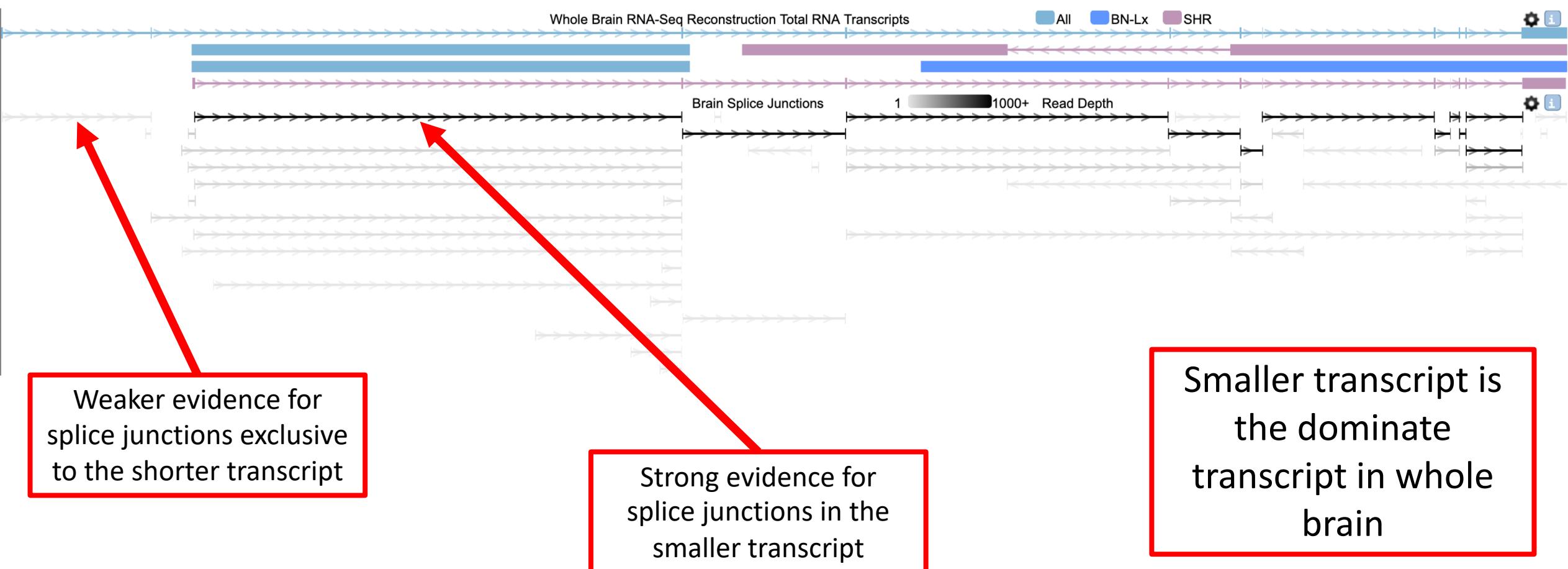


Order	Track Name	Organism	Edit
1	Reference Genomic Sequence i		   
2	Brain Total-RNA Transcriptome i		   
3	Brain Splice Junction Support i	Rat only	   
4	Brain Illumina PolyA+ RNA i	Rat only	   
5	Brain Illumina rRNA-depleted Total-RNA i	Rat only	   
6	Brain Transcriptome Small RNA i	Rat only	   
7	Brain Illumina small RNA i	Rat only	   
8	Ensembl Protein Coding Genes i		   

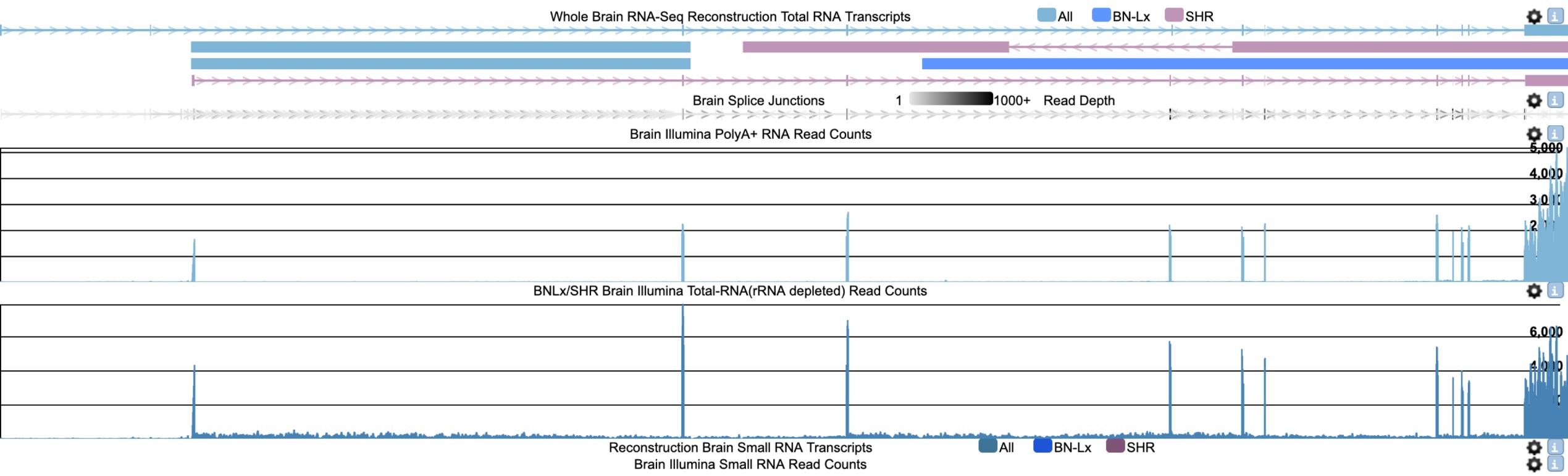
Observations about Transcriptome Reconstruction



Observations About Splice Junctions



Observations About Read Coverage



Peak patterns are similar between polyA-selected RNA sequencing and total RNA sequencing so this transcript is likely polyadenylated.

Peaks for the shorter transcript are well above background noise.

The exon that is exclusive to the longer transcripts is undetectable visually in the presence of the shorter transcript

Examine Brain RNA-Seq Results From Other Inbred Strains

Input Gene of Interest (e.g., Mef2c)

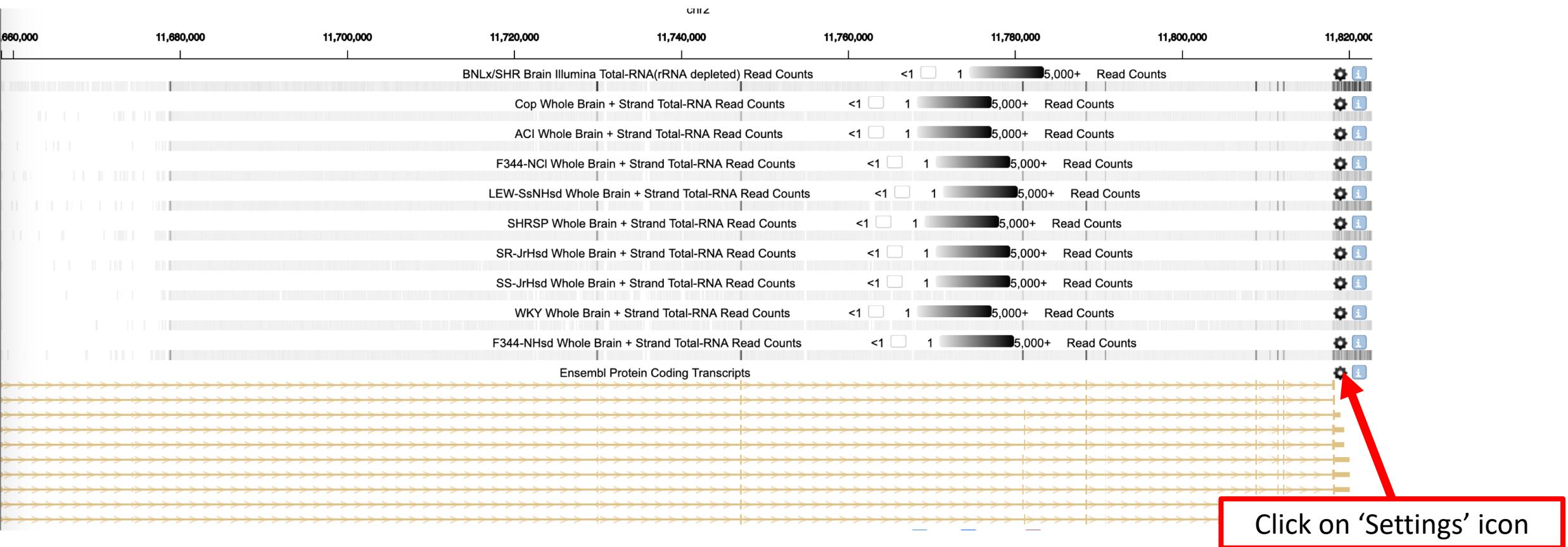
Gene Identifier or Region: Species: ⓘ Initial View: ⏺ Go

Or [Translate Region to Mouse/Rat](#)

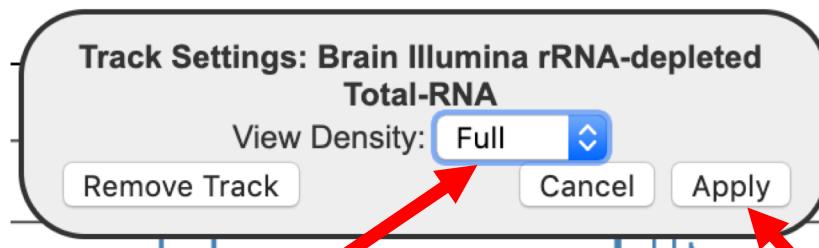
Enter Gene Symbol of Interest

Select 'Inbred Brain Read Depth (Predefined) (Rat Only)' from the drop-down menu for initial view

Extend Read Coverage to 'Full'

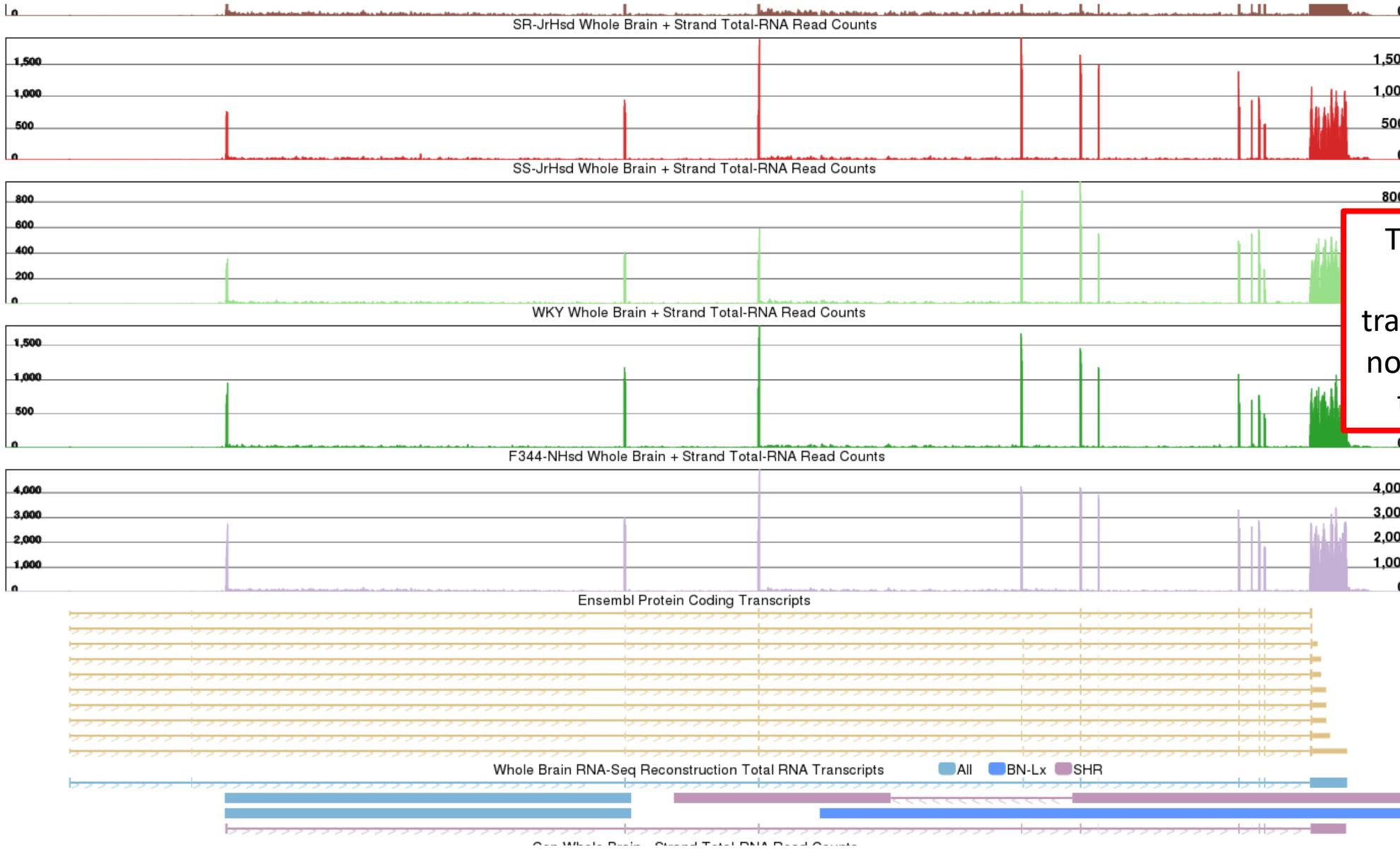


Setting the view to 'Full'



1. Select 'Full' from the drop down menu for 'View Density'

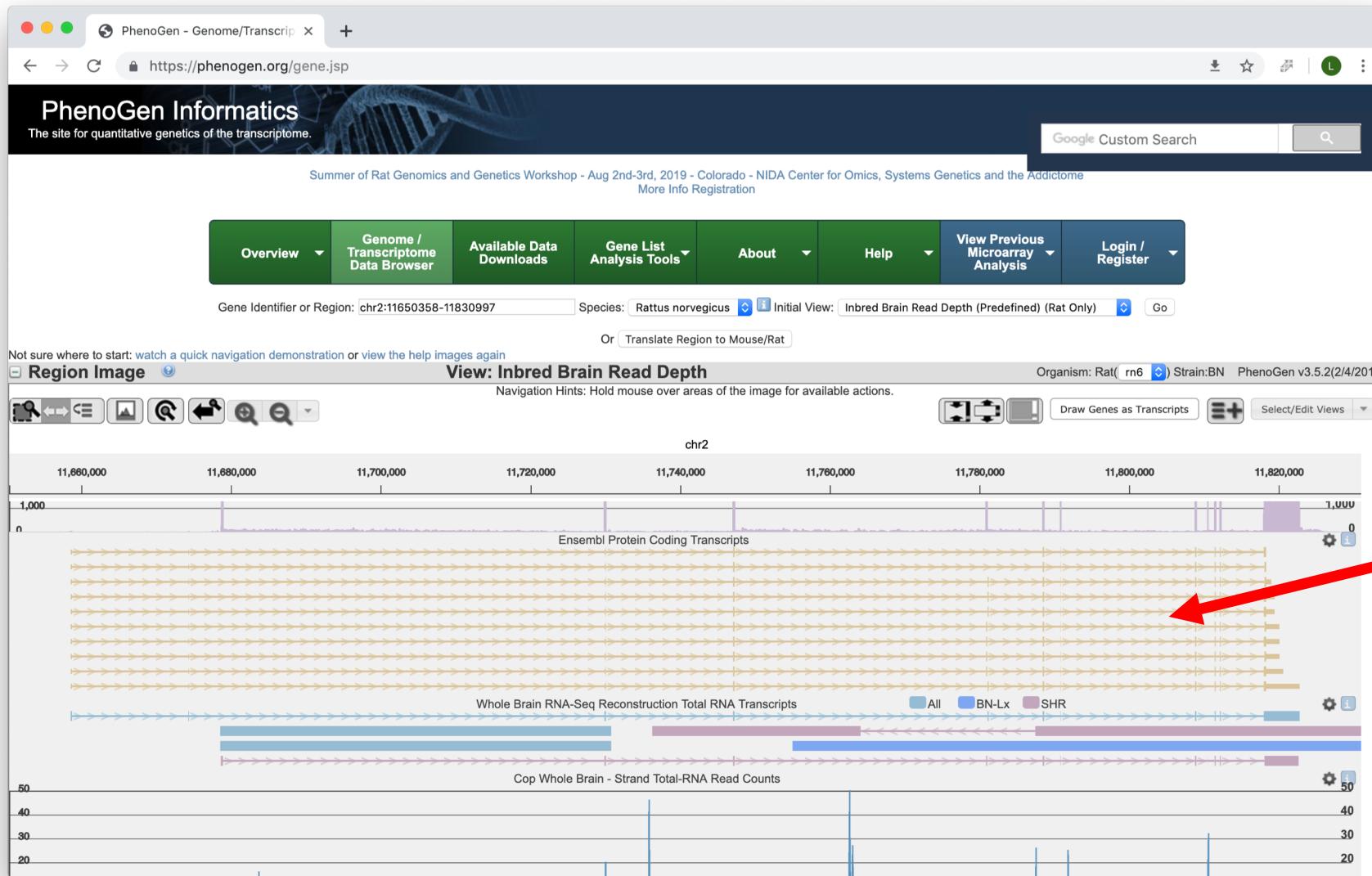
2. Click 'Apply'



The 4 example strains shown have similar transcript structures, i.e., not genetic difference in transcript structure.

Gene Details

How to retrieve Gene Details



Click anywhere in the
Ensembl Protein Coding
Transcript Track



Gene Details

Gene Details ⓘ

Gene eQTLs ⓘ

Expression Data ⓘ

miRNA Targeting Gene(multiMiR) ⓘ

WGCNA ⓘ

Selected Feature Summary

Gene Symbol: Mef2c
Phenogen ID: ENSRNOG00000033134
Location: chr2: 11,658,568-11,822,787
Strand: +
Description: myocyte enhancer factor 2C
Links:
[ENSRNOG00000033134](#)
All Organisms: [NCBI](#) | [UniProt](#)
Rat: [NCBI](#) | [UniProt](#) | [RGD](#)

Links to other databases

Exonic Variants

Exonic Variants:

Common: 0 (SNPs) / 0(Insertions/Deletions)
BN-Lx/CubPrin: 0 (SNPs) / 1(Insertions/Deletions)
SHR/OlaPrin: 8 (SNPs) / 3 (Insertions/Deletions)
SHR/NCrIPrin: 0 (SNPs) / 0 (Insertions/Deletions)
F344: 0 (SNPs) / 0 (Insertions/Deletions)

Transcripts:

[ENSRNOT00000076408](#)
[ENSRNOT00000076416](#)
[ENSRNOT00000076992](#)
[ENSRNOT00000075931](#)
[ENSRNOT00000076136](#)
[ENSRNOT00000076481](#)
[ENSRNOT00000076084](#)
[ENSRNOT00000076230](#)
[ENSRNOT00000076710](#)
[ENSRNOT00000076239](#)

Ensembl-annotated transcripts of this gene

Transcripts derived from the reconstruction and how they compare to annotated transcripts

Gene eQTL

Gene Details ⓘ Gene eQTLs ⓘ Expression Data ⓘ WGCNA ⓘ

Selected Feature Summary

eQTL Image Options ⓘ

Data Source: RNA-Seq ⓘ P-value Threshold for Highlighting: 0.01 ⓘ

Tissues: Include at least one tissue. ⓘ

Excluded	Included
	Brain Liver

Chromosomes: (chr2 must be included) ⓘ

Excluded	Included
	Chr 1 Chr 2 Chr 3 Chr 4 Chr 5 Chr 6

Click to run Circos

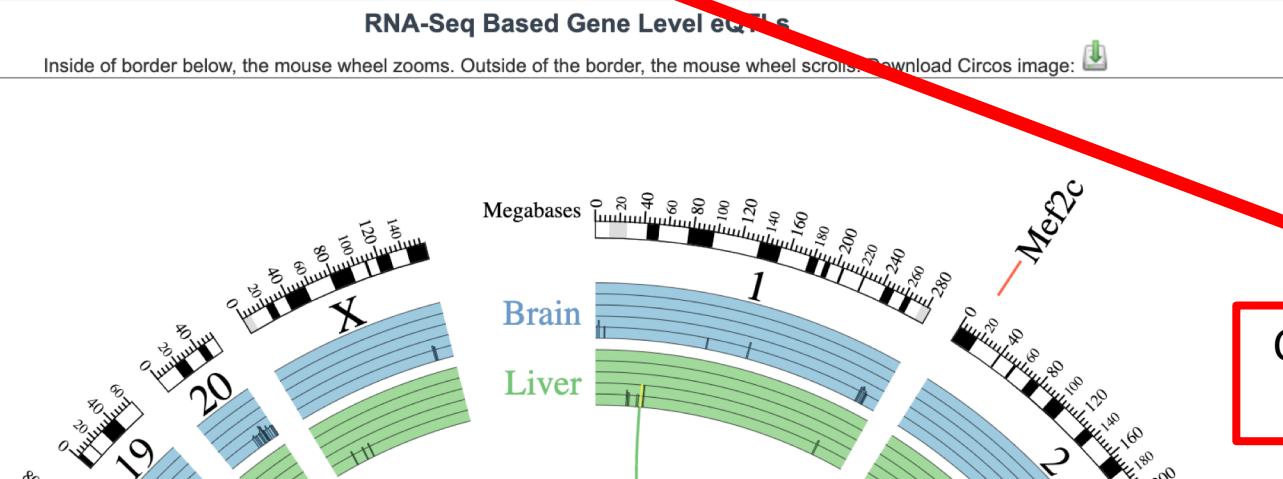
Circos: an Information Aesthetic for Comparative Genomics.

Choose between microarray data or RNA-Seq data

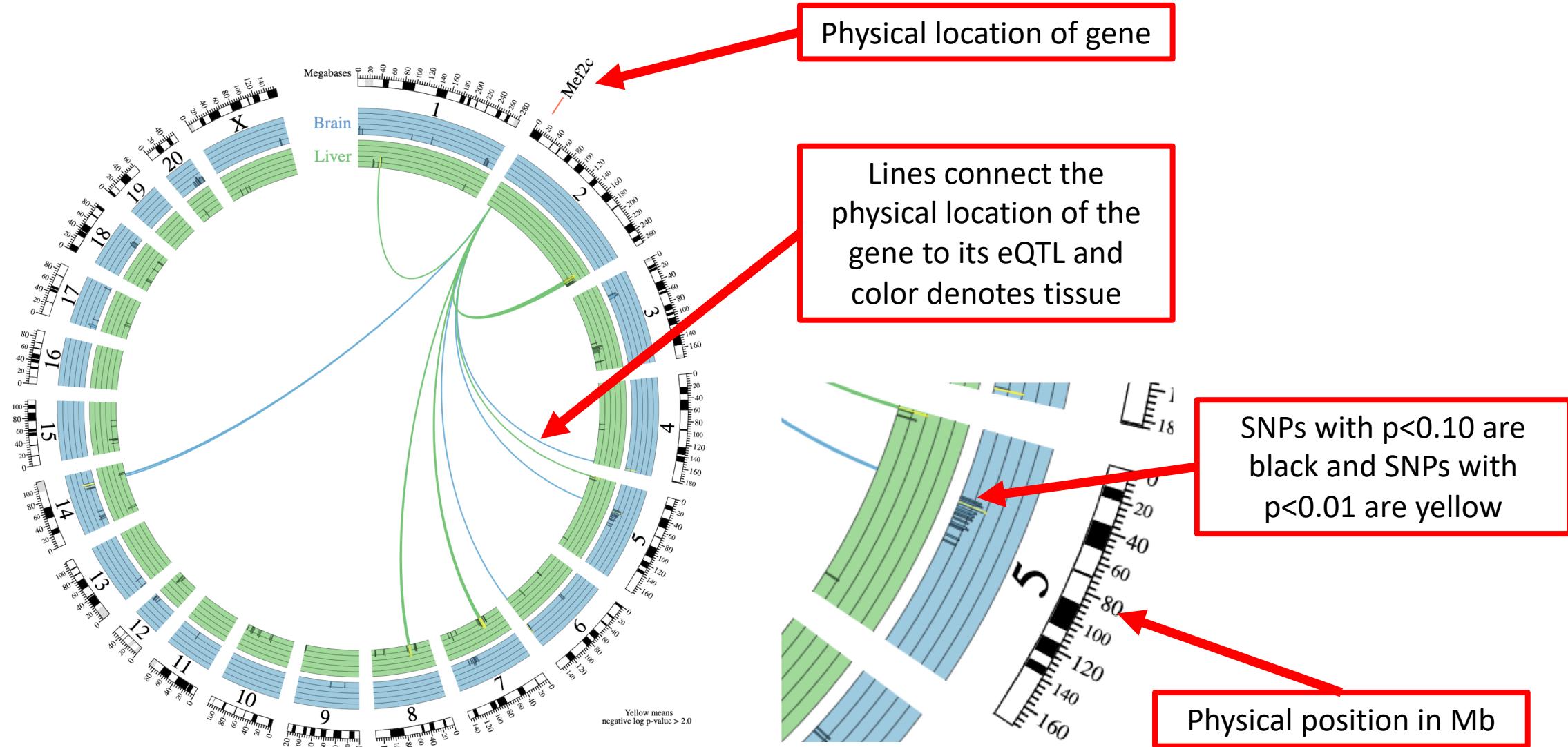
Choose loci to highlight based on p-value for association with expression

Choose which chromosomes to include

Choose which tissues to include

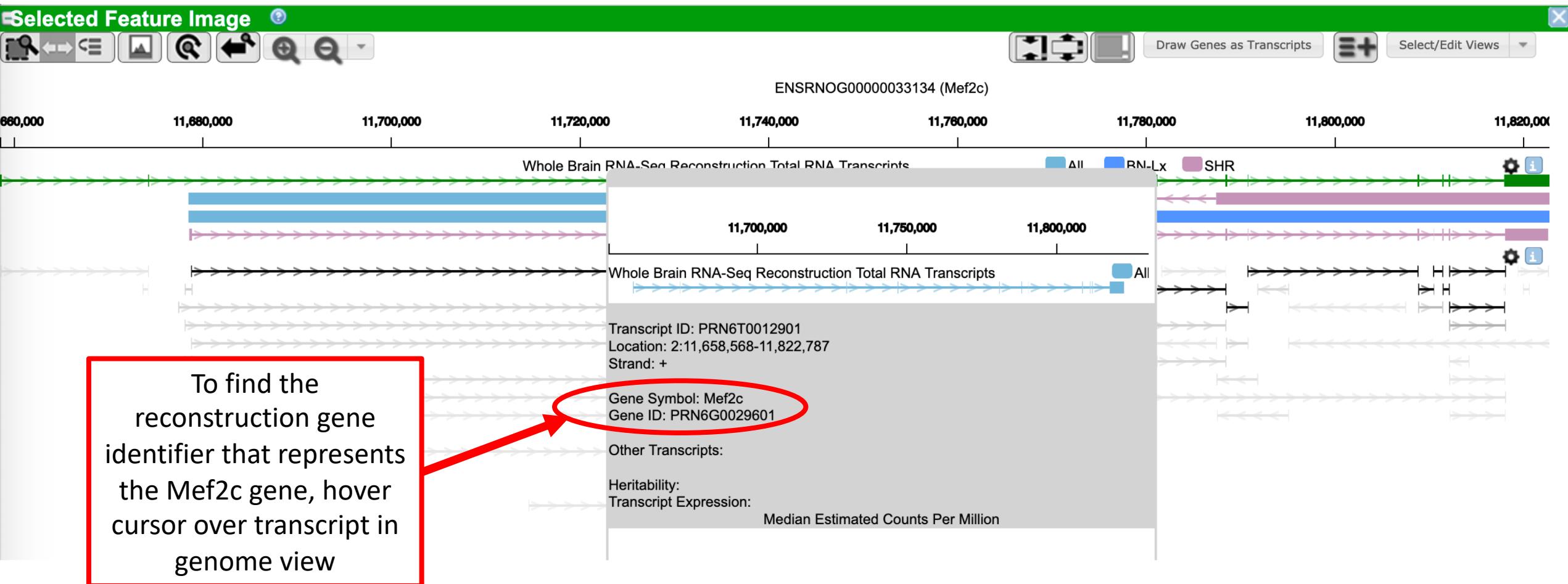


Circos Plot with Gene eQTL



Expression Data

Results Show All Overlapping Reconstruction Genes



Whole Brain Long RNA-Seq Expression



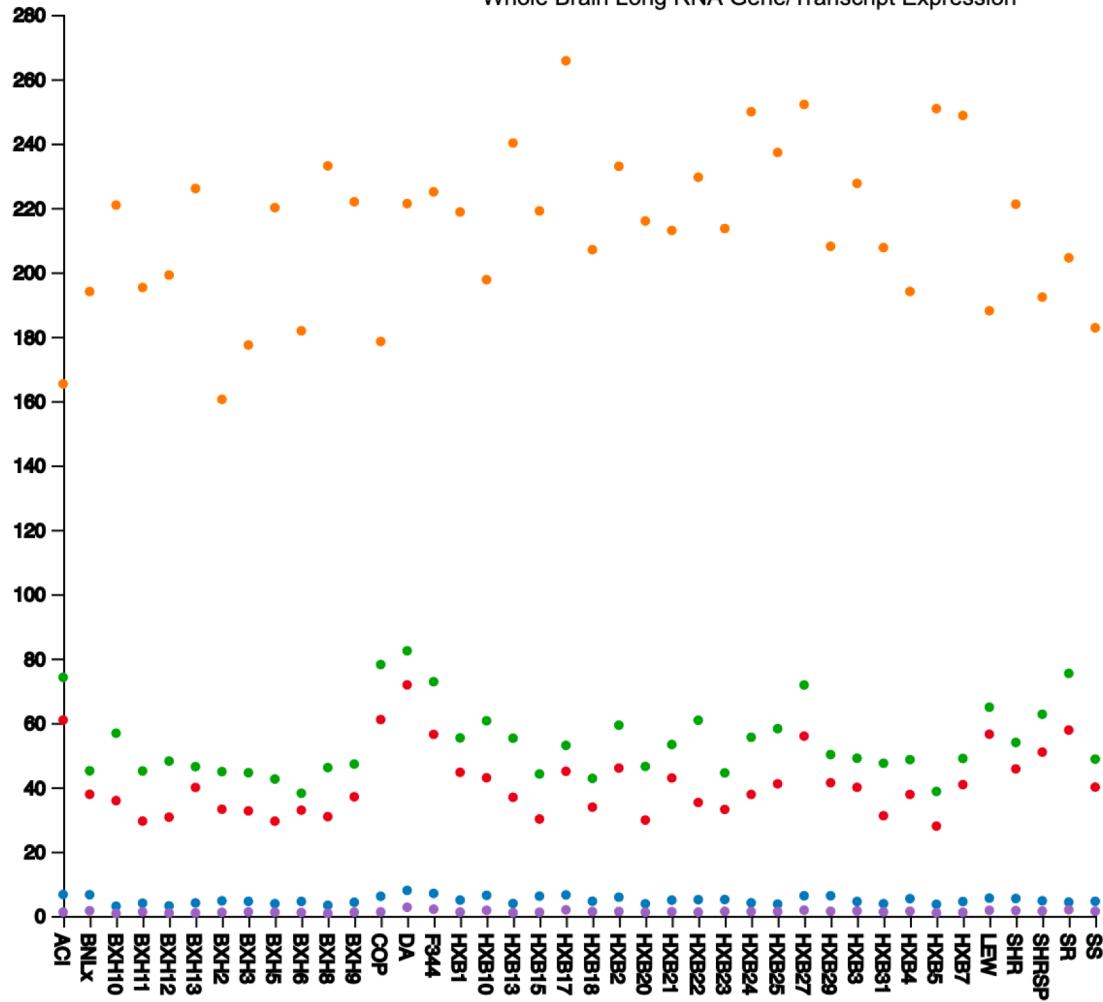
Display: Genes Transcripts Heritability

- PRN6G0029655
- PRN6G0029665
- PRN6G0029602

- PRN6G0029601
- PRN6G0029603

The orange dots represent the Mef2c gene

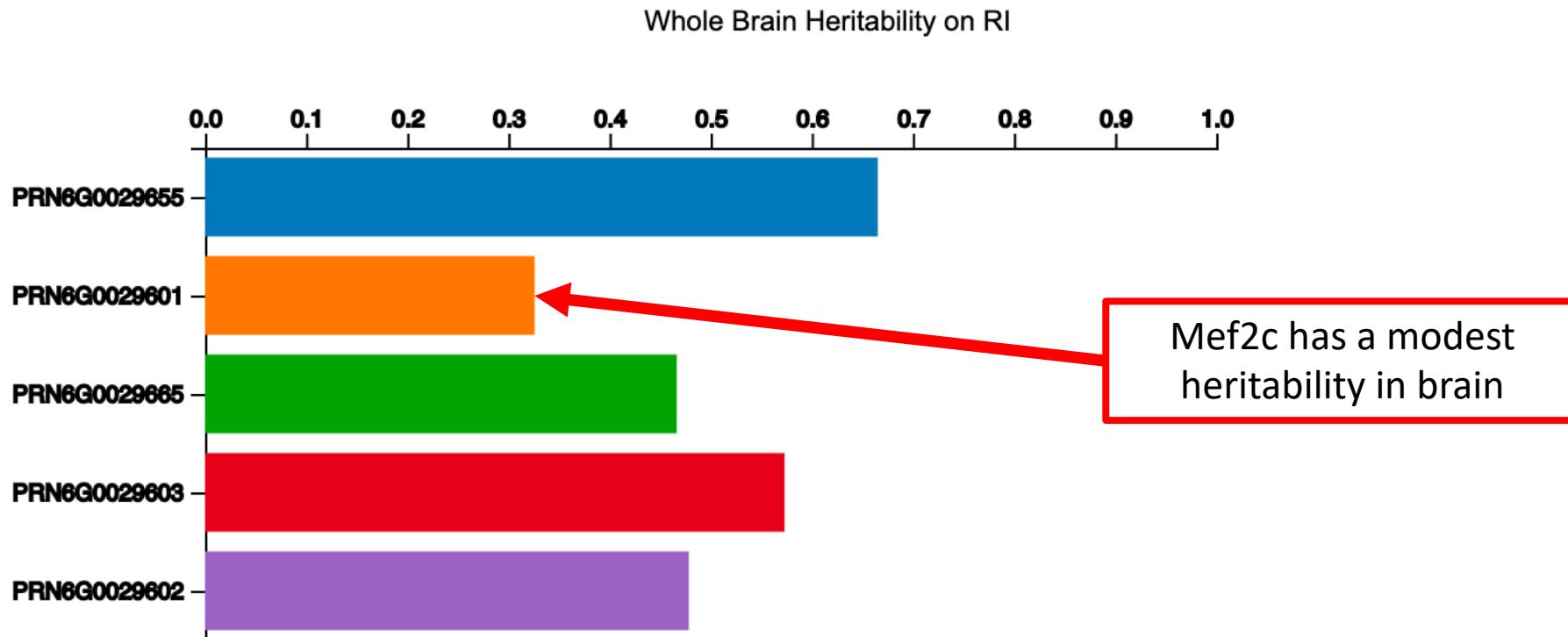
Whole Brain Long RNA Gene/Transcript Expression

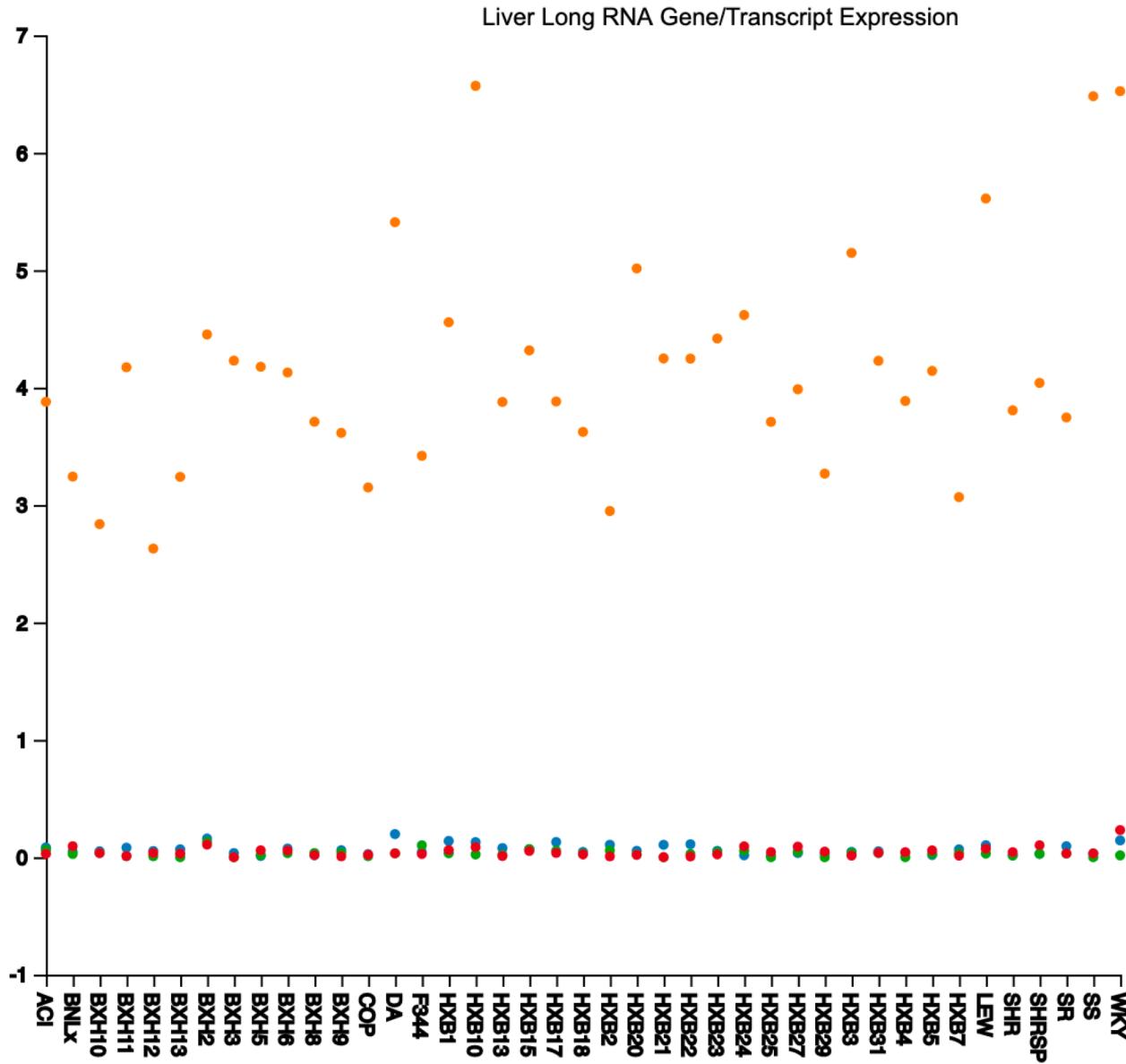


In whole brain, Mef2c is highly expressed with about 200 reads per million mapped reads

Both the highest strain and the lowest strain are from the HXB/BXH recombinant inbred panel
Minimum ~ 160 CPM
Maximum ~ 270 CPM

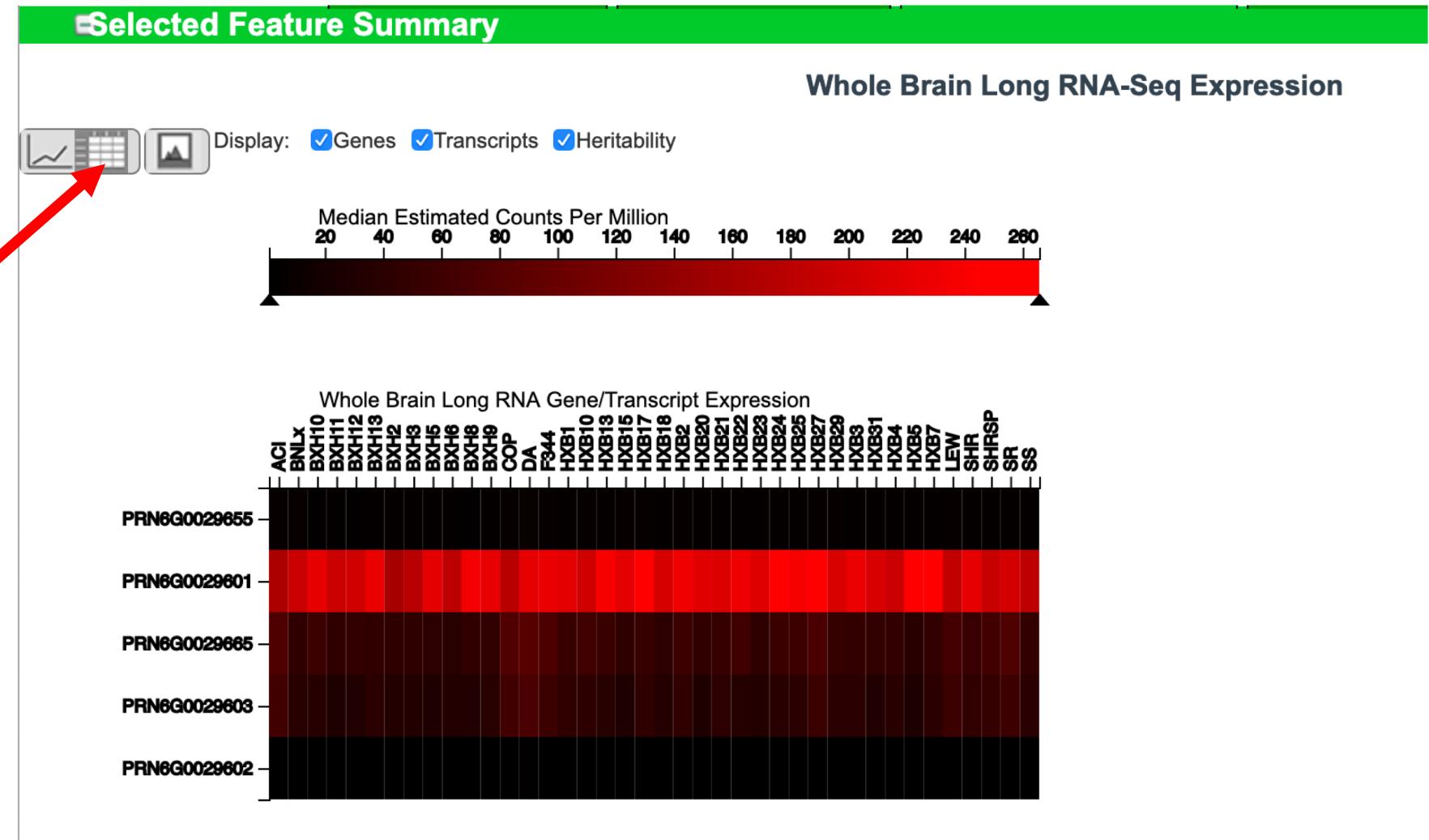
Heritability Whole Brain RNA Expression in the HXB/BXH Panel





In liver, expression levels
of Mef2c are much lower
than in brain (4 CPM)
compared to 220 CPM)

Switch from dotplot to heatmap



miRNA Targeting Gene (multiMir)

List of miRNAs with binding sites on Mef2c

Include both predicted and validated

Details about prediction thresholds used

Selected Feature Summary

multiMiR Options

Validation Level: All

Predicted Cutoff Type: Top percentage of miRNA targets

Predicted Cutoff: Top 20 % of miRNAs

Target Information

Gene Symbol: Mef2c

Gene Entrez ID: 499497

Gene Ensembl ID: ENSRNOG00000033134

multiMiR Results

Search:

Mature miRNA Accession (click for miRBase)	Mature miRNA ID (click to view details)	Validated			Predicted								Total		
		miRecords	miRTarBase	TarBase	DIANA-microT-CDS	EIMMo	MicroCosm	miRanda	miRDB	PicTar	PITA	TargetScan	Total Validated	Total Predicted	Total All
MIMAT0000611	rno-miR-135b-5p	-	-	-	-	-	-	-	X	-	-	-	0	1	1
MIMAT0000781	rno-miR-9a-5p	-	-	-	-	-	-	-	X	-	-	-	0	1	1
MIMAT0000809	rno-miR-30a-3p	-	-	-	-	-	X	-	-	-	-	-	0	1	1
MIMAT0000841	rno-miR-135a-5p	-	-	-	-	-	-	-	X	-	-	-	0	1	1
MIMAT0000860	rno-miR-183-5p	-	-	-	-	-	-	-	X	-	-	-	0	1	1
MIMAT0000876	rno-miR-203a-3p	-	-	-	-	-	-	-	X	-	-	-	0	1	1
MIMAT0003194	rno-miR-376c-3p	-	-	-	-	-	-	-	X	-	-	-	0	1	1
MIMAT0004639	rno-miR-325-3p	-	-	-	-	-	-	-	X	-	-	-	0	1	1
MIMAT0004722	rno-miR-30d-3p	-	-	-	-	-	X	-	-	-	-	-	0	1	1
MIMAT0004743	rno-miR-300-5p	-	-	-	-	-	X	-	-	-	-	-	0	1	1

Co-expression modules (WGCNA)

Weighted Gene Co-expression Network Analysis (WGCNA)

Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 4

Why not just correlation?

1. Relationship between genes better described using a **scale-free network**
2. We can calculate a more **robust** measure of co-expression by including a measure of how many “friends” two genes have in common (indirect relationships)

WGCNA Tab

Co-expression networks are derived from the HXB/BXH Recombinant Inbred panel, but you can choose the tissue and the source of RNA expression levels

Gene Details ⓘ Gene eQTLs ⓘ Expression Data ⓘ miRNA Targeting Gene(multiMiR) ⓘ WGCNA ⓘ

Selected Feature Summary

Data Source: RNA-Seq Tissue: Whole Brain

Navigation Hints: Hold mouse over areas of the image for available actions.

Views:

- Module ⓘ
- Eigengene eQTL ⓘ
- Gene Ontology ⓘ
- miRNA Targets ⓘ
- Meta Module (nearest neighbors) ⓘ

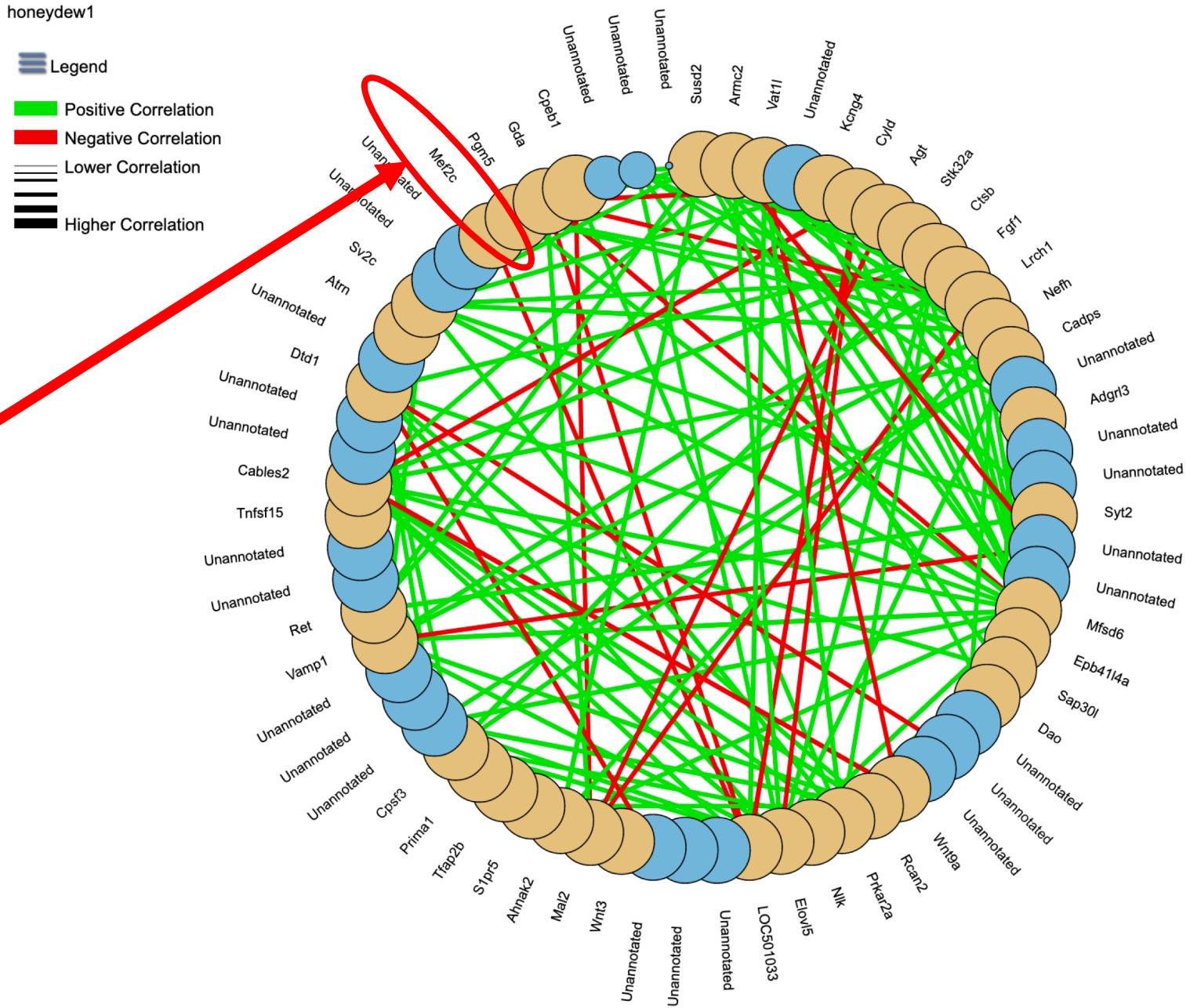
honeydew1

Mef2c is in the honeydew1 module. Click on the circle to get more details about the this module

The ‘Module’ view displays the correlations between genes

Co-expression module

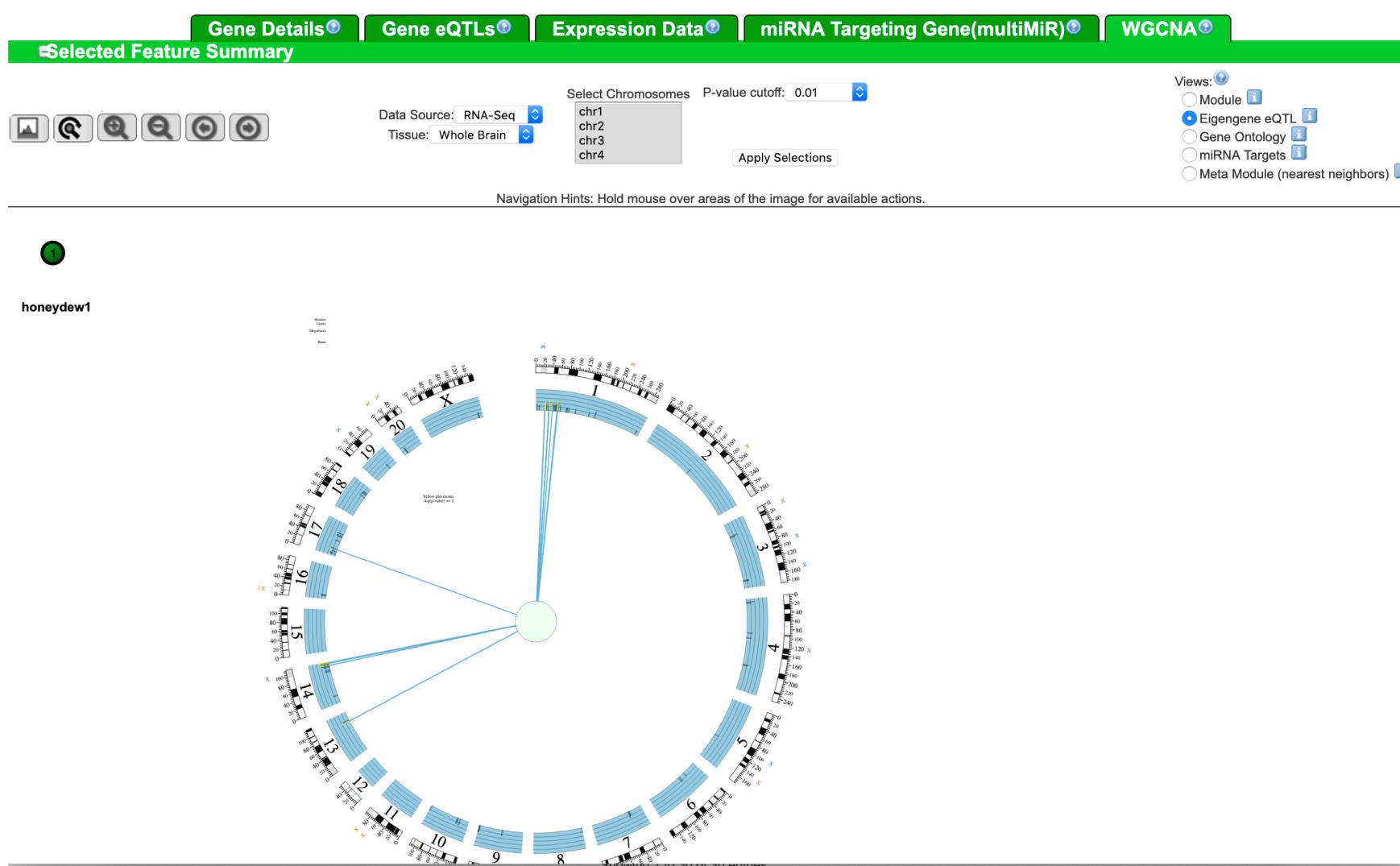
Mef2c is not one of the most connected genes in the model. It is negatively correlated with LOC501033 and positively correlated with Susd2 (hub gene, i.e., most connected gene)



Information on individual genes is given in the table below the module graphic

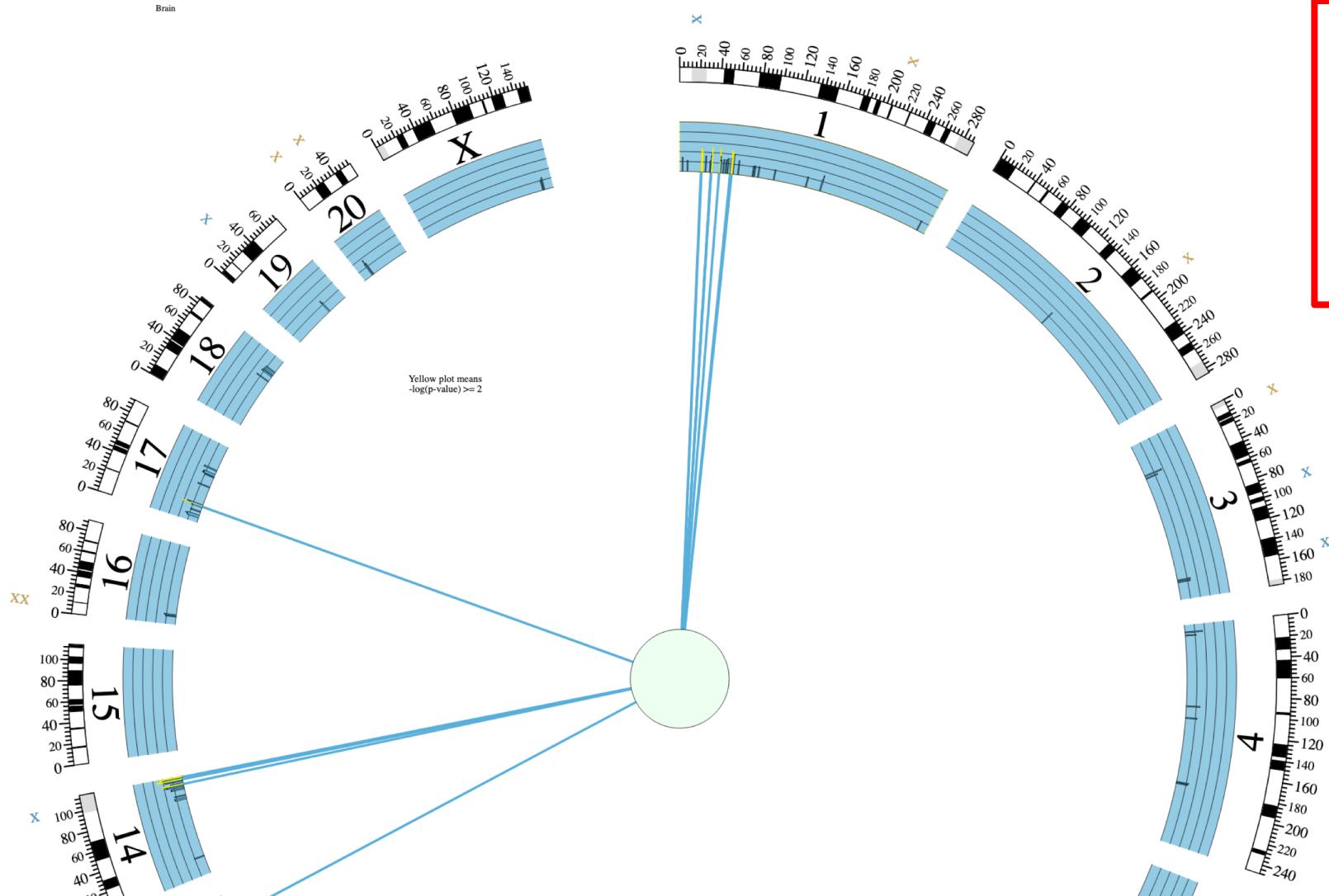
Transcripts in honeydew1 Module					Export CSV
Gene Symbol	Gene ID	Transcript	Probe Sets	Link Total	▲
Susd2	ENSRNOG00000033389	ENSRNOG00000033389	0	37.68	
Armc2	ENSRNOG00000026217	ENSRNOG00000026217	0	37.48	
Vat1l	ENSRNOG00000011989	ENSRNOG00000011989	0	34.71	
Unannotated	PRN6G0244905	PRN6G0244905	0	34.68	
Kcng4	ENSRNOG00000015746	ENSRNOG00000015746	0	34.42	
Cyld	ENSRNOG00000014048	ENSRNOG00000014048	0	33.80	
Agt	ENSRNOG00000018445	ENSRNOG00000018445	0	31.91	
Stk32a	ENSRNOG00000027066	ENSRNOG00000027066	0	31.08	
Ctsb	ENSRNOG00000010331	ENSRNOG00000010331	0	31.03	
Fgf1	ENSRNOG00000013867	ENSRNOG00000013867	0	30.86	
Lrch1	ENSRNOG00000009412	ENSRNOG00000009412	0	30.79	

Eigengene eQTL



The module eigengene is a summary measure of the expression variation across the genes within the module

Eigengene eQTL Circos Plot



Similar to the previous eQTL circos plot, the vertical bars represent the negative log base 10 p-value at an individual locus.

The 'X' along the outside of the circle represent the physical location of different genes in the module

Gene Ontology

Gene Details ⓘ | Gene eQTLs ⓘ | Expression Data ⓘ | miRNA Targeting Gene(multiMiR) ⓘ | WGCNA ⓘ

Selected Feature Summary

Data Source: RNA-Seq ⏺ GO Domain: Molecular Function ⏺
Tissue: Whole Brain ⏺ Number of Levels to Display: 3

Views: ⓘ

- Module ⓘ
- Eigengene eQTL ⓘ
- Gene Ontology ⓘ **Selected**
- miRNA Targets ⓘ
- Meta Module (nearest neighbors) ⓘ

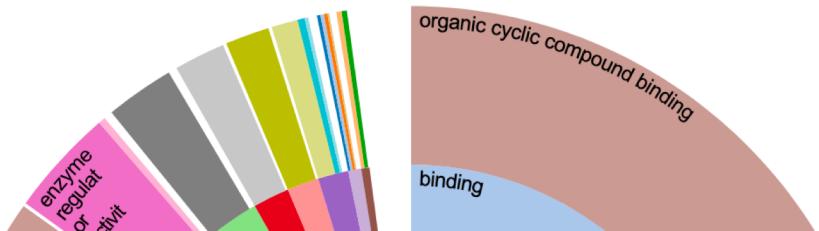
Click to select this module and see additional details.

1 honeydew1

35 genes in module with GO annotation

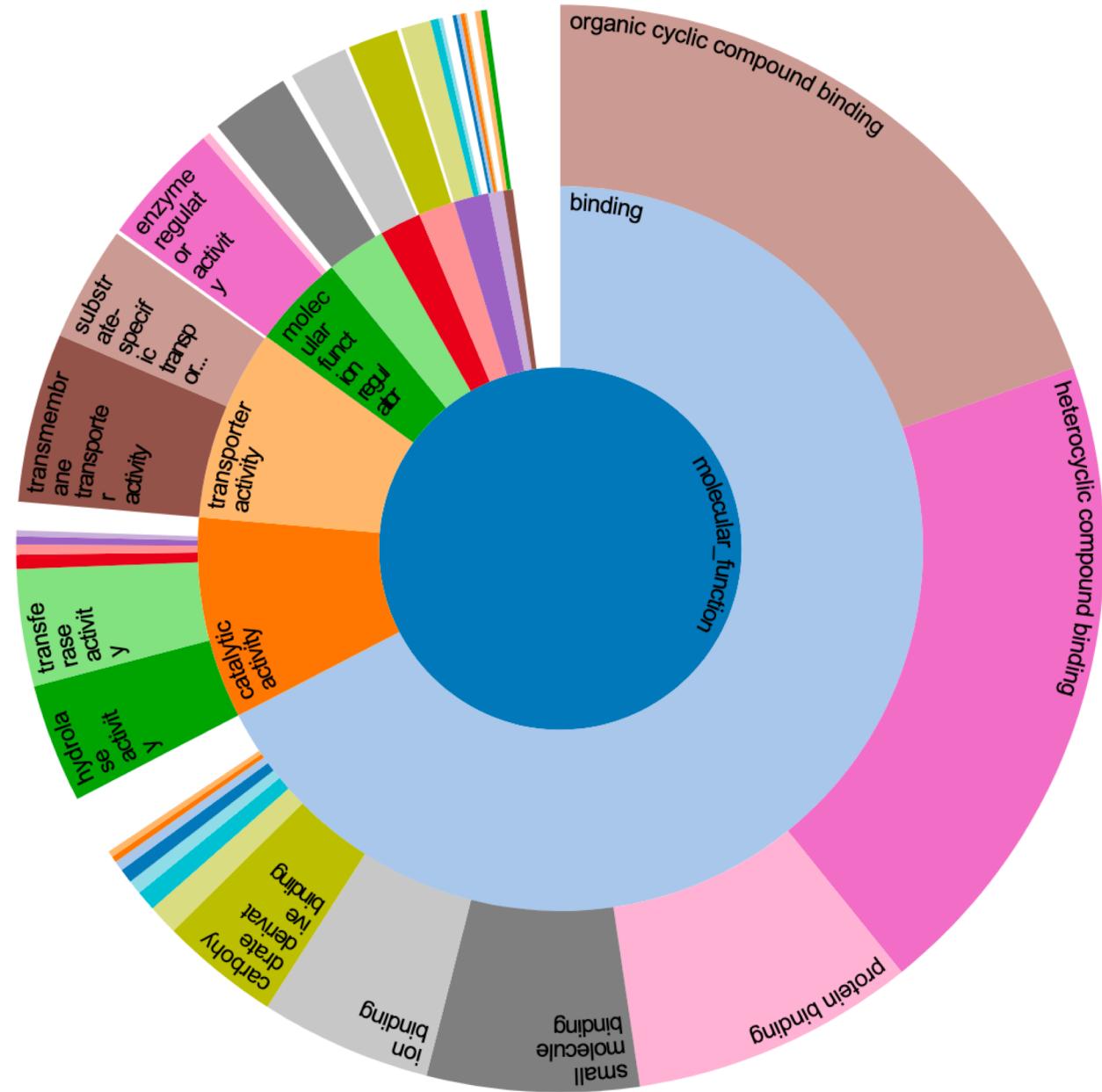
You can change the GO domain and the number of levels that are shown at a time in the sunburst plot below

This summarizes the genes within the module based on shared gene ontology terms



GO Term Sunburst Plot

Each color represents a different GO term. The GO term underneath it is a parent term and each term outside a child term.



honeydew1

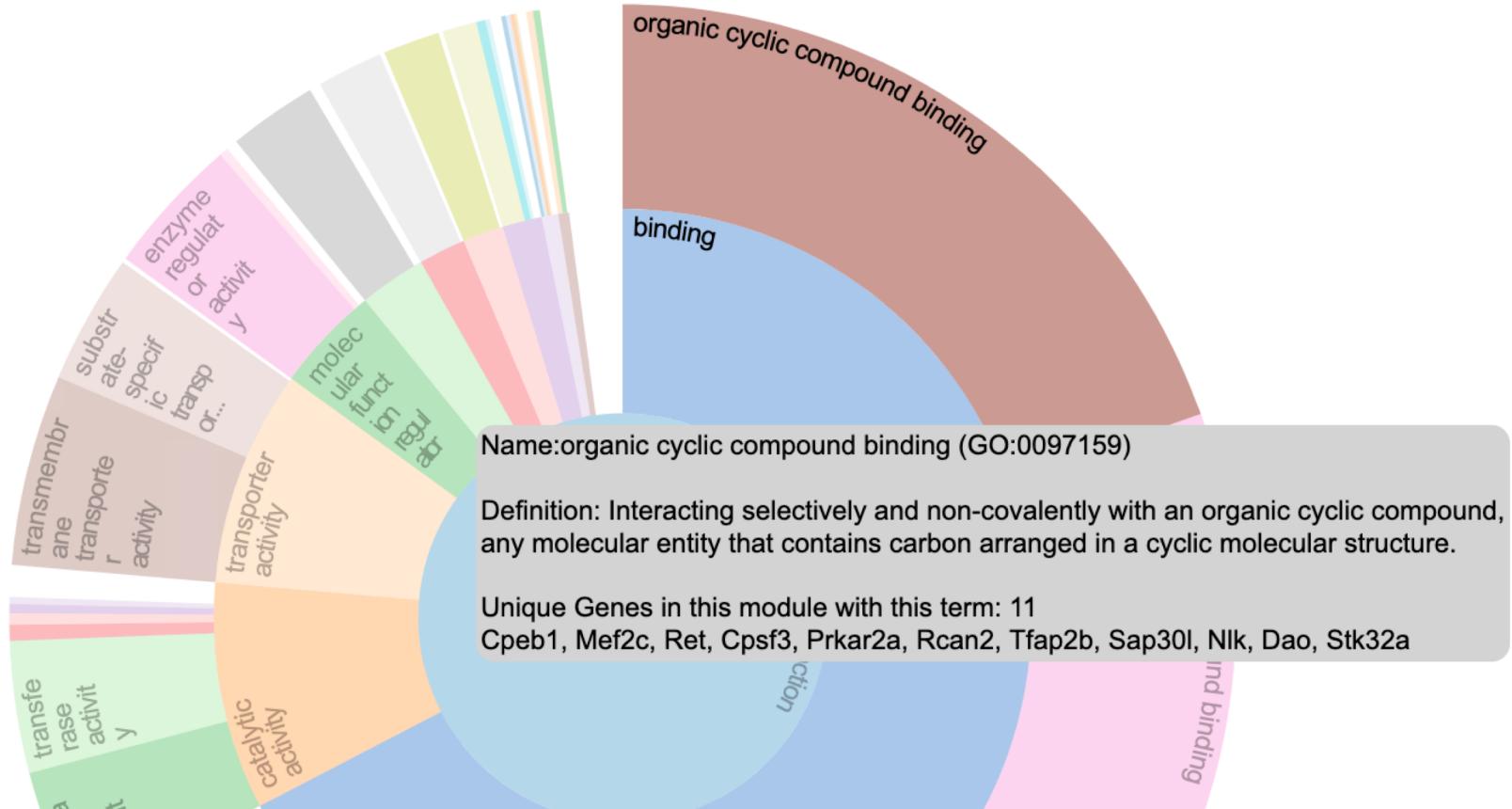
binding

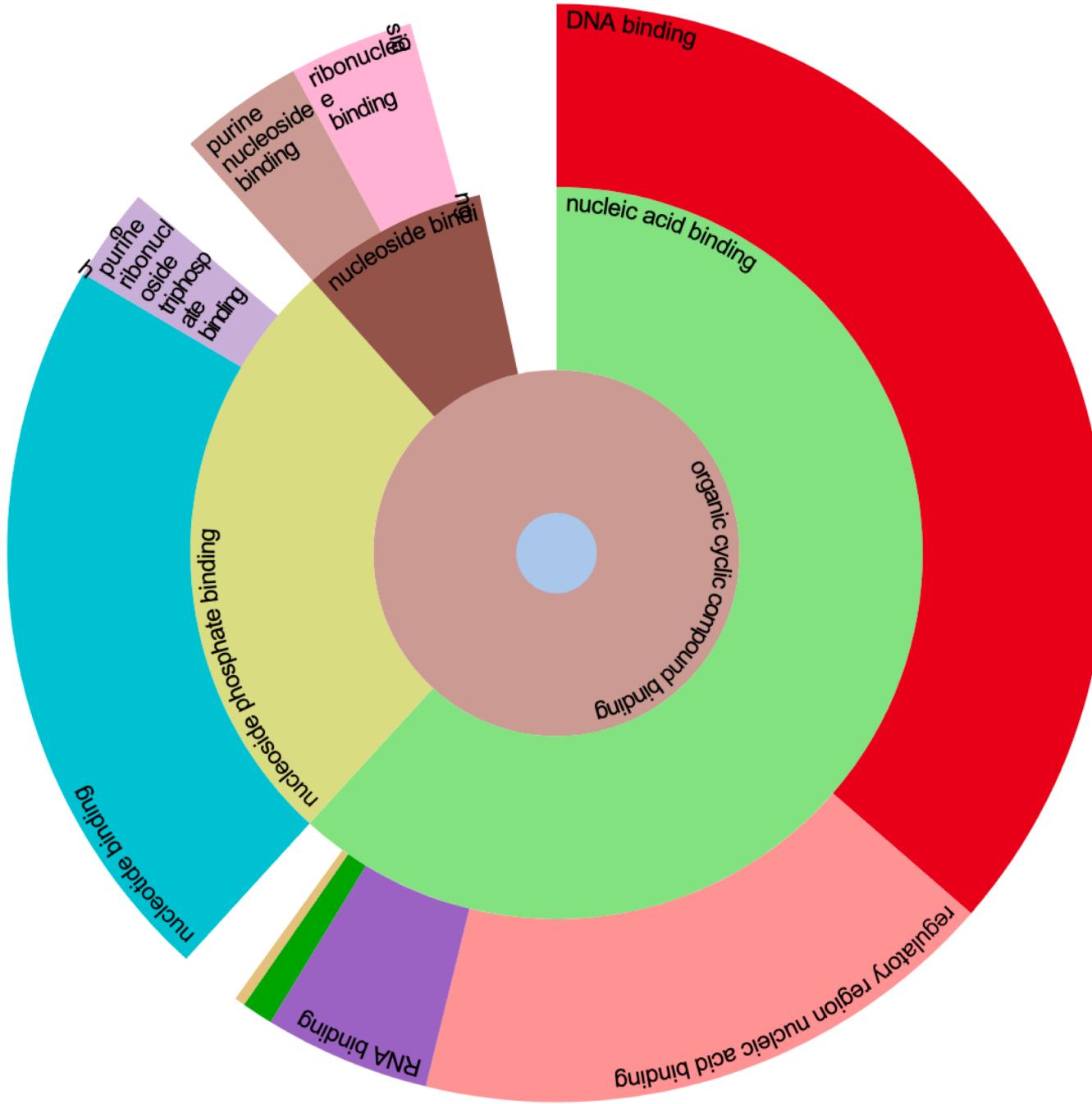
organic cyclic compound binding

11 genes (31.4%)

35 genes in module with GO annotation

Hover over any GO term
to see definition and
genes from the module
annotated to this GO term





You can also click on any term and see the next three layers of child terms that include genes from the module